

PENERAPAN *TEXT MINING* DAN ANALISIS SENTIMEN UNTUK REKOMENDASI HOTEL BERDASARKAN *ONLINE REVIEW*

Tugas Akhir

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat
Sarjana Komputer**



Dibuat Oleh:

MARGARETA OKTAVIANA SHINTAPURI

16 07 08707

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2020**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

PENERAPAN TEXT MINING DAN ANALISIS SENTIMEN UNTUK REKOMENDASI HOTEL
MENGUNAKAN REVIEW ONLINE

yang disusun oleh

MARGARETA OKTAVIANA SHINTAPURI

160708707

dinyatakan telah memenuhi syarat pada tanggal 11 Agustus 2020

		Keterangan
Dosen Pembimbing 1	: Prof. Ir. A. Djoko Budiyanto, M.Eng., Ph.D.	Telah menyetujui
Dosen Pembimbing 2	: Ike Devi Sulistyaningtyas, S.Sos., M. Si.	Telah menyetujui
Tim Penguji		
Penguji 1	: Prof. Ir. A. Djoko Budiyanto, M.Eng., Ph.D.	Telah menyetujui
Penguji 2	: Dr. Andi Wahyu Rahardjo, BSEE., MSSE	Telah menyetujui
Penguji 3	: Prof. Ir. Suyoto, MSc., PhD	Telah menyetujui

Yogyakarta, 11 Agustus 2020

Universitas Atma Jaya Yogyakarta

Fakultas Teknologi Industri

Dekan

ttd

Dr. A. Teguh Siswanto, M.Sc

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Margareta Oktaviana Shintapuri
NPM : 160708707
Program Studi : Teknik Informatika
Fakultas : Teknologi Industri
Judul Penelitian : Penerapan *Text Mining* dan Analisis Sentimen
untuk Rekomendasi Hotel berdasarkan *Online Review*

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Surabaya, 3 Agustus 2020

Yang menyatakan,

Margareta Oktaviana Shintapuri

16 07 08707

HALAMAN PERSEMBAHAN

Semua akan indah pada waktu-Nya



KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “Penerapan *Text Mining* dan Analisis Sentimen untuk Rekomendasi Hotel berdasarkan *Online Review*” ini dengan baik.

Penulisan tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mencapai derajat sarjana komputer dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapatkan bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu membimbing dalam iman-Nya, memberikan berkat-Nya, dan menyertai penulis selalu.
2. Bapak Dr. A. Teguh Siswantoro, M.Sc., selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Bapak Martinus Maslim, S.T., M.T., selaku Kepala Program Studi Teknik Informatika, Universitas Atma Jaya Yogyakarta
4. Bapak Prof. Ir. A. Djoko Budiyo SHR, M.Eng., Ph.D, selaku dosen pembimbing I yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
5. Ibu Ike Devi Sulistyningtyas, M. Si., selaku dosen pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
6. Kedua orang tua yang selalu mendoakan dan memberi dukungan baik moral dan material kepada penulis untuk selalu berusaha mencapai hasil yang terbaik.

7. Para sahabat dan teman-teman yang selalu mendoakan dan menyemangati penulis selama berproses di jenjang kuliah.
8. Semua pihak yang tidak dapat disebutkan satu per satu yang selalu memberikan doa, dukungan, dan semangat kepada penulis.

Demikian laporan tugas akhir ini dibuat, dan penulis mengucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Surabaya, 3 Agustus 2020

Margareta Oktaviana Shintapuri

16 07 08707



DAFTAR ISI

HALAMAN PENGESAHAN.....	i
LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH.....	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
INTISARI.....	xi
BAB I. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian.....	5
1.5 Metode Penelitian.....	5
1.6 Sistematika Penelitian.....	6
BAB II. TINJAUAN PUSTAKA.....	8
BAB III. LANDASAN TEORI.....	18
BAB IV. ANALISIS DAN PERANCANGAN EKSPERIMEN	29
4.1 Deskripsi Problem	29
4.2 Analisis Kebutuhan Eksperimen	29
4.3 Perancangan Eksperimen.....	30
BAB V. HASIL EKSPERIMEN DAN PEMBAHASAN	52
5.1 Deskripsi Eksperimen.....	52
5.2 Hasil Eksperimen.....	53
5.3 Pembahasan Eksperimen	68
BAB VI. PENUTUP	72
6.1 Kesimpulan.....	72

6.2 Saran	72
DAFTAR PUSTAKA	74



DAFTAR GAMBAR

Gambar 3.1 Proses CRM menurut Payne	23
Gambar 3.2 Model dari delapan layanan	25
Gambar 4.1 Mengambil <i>review</i> dari situs <i>TripAdvisor.com</i>	30
Gambar 4.2 Hasil pengambilan <i>review</i> dari situs <i>TripAdvisor.com</i>	30
Gambar 4.3 Contoh Kamus Stopwords.....	34
Gambar 4.4 Hasil dari pengolahan menggunakan metode <i>sentistrength</i>	40
Gambar 4.5 Hasil proses menggunakan kamus rekomendasi	44
Gambar 4.6 Hasil proses menggunakan perhitungan nilai untuk rekomendasi secara implisit	45
Gambar 4.7 Hasil perhitungan frekuensi menggunakan <i>n-gram</i>	46
Gambar 4.8 Hasil proses pengkodean delapan mode layanan	50
Gambar 5.1 Diagram presentase hasil analisis sentimen Hotel Eastparc.....	53
Gambar 5.2 Diagram presentase hasil rekomendasi secara <i>eksplisit</i> Hotel Eastparc.....	54
Gambar 5.3 Diagram presentase hasil rekomendasi secara <i>implisit</i> Hotel Eastparc.....	55
Gambar 5.4 Diagram presentase hasil analisis sentimen Hotel The Phoenix	58
Gambar 5.5 Diagram presentase hasil rekomendasi secara <i>eksplisit</i> Hotel The Phoenix	59
Gambar 5.6 Diagram presentase hasil rekomendasi secara <i>implisit</i> Hotel The Phoenix	60
Gambar 5.7 Diagram presentase hasil analisis sentimen Hotel Tentrem.....	63
Gambar 5.8 Diagram presentase hasil rekomendasi secara <i>eksplisit</i> Hotel Tentrem	64
Gambar 5.9 Diagram presentase hasil rekomendasi secara <i>implisit</i> Hotel Tentrem	65

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian.....	16
Tabel 4.1 Contoh data mentah	31
Tabel 4.2 Contoh data bersih hasil dari proses <i>pre-processing</i>	34
Tabel 4.3 Contoh hasil dari pengolahan menggunakan metode <i>sentistrength</i>	37
Tabel 4.4 Contoh kamus rekomendasi	41
Tabel 4.5 Contoh hasil dari pengolahan menggunakan kamus rekomendasi	42
Tabel 4.6 Kode delapan mode layanan	47
Tabel 4.7 Contoh kamus delapan mode layanan.....	47
Tabel 4.8 Contoh hasil pengkodean delapan mode layanan	48
Tabel 5.1 15 kata yang sering muncul pada <i>review online</i> Hotel Eastparc.....	56
Tabel 5.2 Hasil identifikasi delapan mode layanan Hotel Eastparc.....	57
Tabel 5.3 15 kata yang sering muncul pada <i>review online</i> Hotel The Phoenix	61
Tabel 5.4 Hasil identifikasi delapan mode layanan Hotel The Phoenix	62
Tabel 5.5 15 kata yang sering muncul pada <i>review online</i> Hotel Tentrem.....	66
Tabel 5.6 Hasil identifikasi delapan mode layanan Hotel Tentrem.....	67

INTISARI

PENERAPAN *TEXT MINING* DAN ANALISIS SENTIMEN UNTUK REKOMENDASI HOTEL BERDASARKAN *ONLINE REVIEW*

Intisari

Margareta Oktaviana Shintapuri

16 07 08707

Peningkatan jumlah *review* dari banyak pengalaman membuat wisatawan sulit untuk membaca semua *review* untuk mengambil keputusan yang baik untuk memuaskan diri mereka. Oleh karena itu, pelanggan yang akan membuat rekomendasi layanan secara langsung dengan menggunakan ekspresi seperti "saya tidak merekomendasikan" dapat membantu konsumen dalam proses pengambilan keputusan mereka. Selain itu, pelanggan yang menuliskan pujian berkali-kali juga dapat membantu wisatawan secara tidak langsung. *Review* seperti itu menunjukkan bahwa pelanggan puas sehingga dia akan menyarankan orang lain untuk mencoba.

Metode *text mining* diterapkan pada *review online* yang telah dikumpulkan dari situs *TripAdvisor.com* untuk tiga *review* terbanyak pada hotel yang ada di DI Yogyakarta saat pengambilan data. Data yang dikumpulkan ini di analisis sentimennya sehingga dibagi menjadi kelas positif, negatif, dan netral. Setelah itu, data akan di analisis rekomendasi secara *eksplisit* dan secara *implisit*. Kemudian hasil *review* yang direkomendasikan akan diproses dengan menggunakan *n-gram* untuk melihat kata yang sering keluar yang akan diproses lagi dan menghasilkan delapan mode layanan tambahan pada hotel yang direkomendasikan.

Hasil dari penelitian ini, diharapkan agar wisatawan atau pengunjung hotel mendapatkan pelayanan yang baik serta mendapatkan kepuasan dari pelayanan hotel tersebut dari rekomendasi pelanggan yang pernah menginap di hotel dan meninggalkan ulasan pada situs *TripAdvisor.com*.

Kata Kunci: *text mining*, *rekomendasi eksplisit*, *rekomendasi implisit*, *sentiment strength*

Dosen Pembimbing I : Prof. Ir. Djoko Budiyanto SHR A, M.Eng., Ph.D

Dosen Pembimbing II : Ike Devi Sulistyaningtyas, S.Sos., M.Si.

Jadwal Sidang Tugas Akhir : xxxx

BAB I. PENDAHULUAN

1.1. Latar Belakang

Yogyakarta merupakan salah satu kota di Indonesia yang menjadi destinasi bagi wisatawan nusantara (wisnus) dan wisatawan mancanegara (wisman). Hal tersebut dikarenakan kota Yogyakarta juga merupakan kota perjuangan, pusat pendidikan, dan pusat kebudayaan. Destinasi wisata yang diberikan untuk dari pengunjung dalam dan dari luar kota Yogyakarta terbagi menjadi lima jenis yaitu Kampung Wisata Jogja, Wisata Museum, Wisata Sejarah dan Budaya, Wisata Religi, dan Wisata Kuliner. Pada data statistik pariwisata Daerah Istimewa Yogyakarta tahun 2018, jumlah destinasi wisata yang ada sebanyak 185 destinasi dan jumlah wisatawan ke destinasi tersebut tahun 2018 berjumlah 25.915.686 wisatawan nusantara dan 600.102 wisatawan mancanegara. Jadi, jika dijumlahkan total wisatawan yang datang berjumlah 26.515.788 orang[1].

Sebagai tempat destinasi wisata hotel merupakan salah satu akomodasi terpenting saat ingin berlibur atau sekedar datang pada suatu daerah. Di kota Yogyakarta mempunyai 685 hotel yang dapat dikunjungi oleh wisatawan yang dapat dibagi menjadi hotel bintang dan hotel non bintang. Hotel bintang memiliki 5 jenis lagi yaitu hotel bintang 1 dengan jumlah 13 hotel, hotel bintang 2 dengan jumlah 24 hotel, hotel bintang 3 dengan jumlah 32 hotel, hotel bintang 4 dengan jumlah 18 hotel, dan hotel bintang 5 dengan jumlah 9 hotel. Sedangkan, untuk hotel non bintang berjumlah 589 hotel. Jumlah wisatawan yang menggunakan jasa akomodasi hotel tahun 2018 berjumlah 5.272.719 wisatawan nusantara dan 416.372 wisatawan mancanegara. Jadi, jika dijumlahkan total wisatawan yang datang berjumlah 5.689.019 orang[1].

Dalam membuat suatu bisnis konsultan AS Michael Treacy dan Fred Wiersema menyarankan tiga strategi bisnis yang luas yaitu keunggulan operasional, kepemimpinan produk, dan keintiman pelanggan. Mereka juga mengatakan bahwa keberhasilan pasar biasanya didasarkan pada jenis proposisi nilai yang akan dikerjar oleh perusahaan contohnya seperti biaya

terbaik, produk terbaik, atau solusi terbaik. Tiga strategi bisnis yang mereka sarankan memiliki arti sebagai berikut: keunggulan operasional yang dimaksudkan adalah menyediakan pelanggan dengan produk atau layanan dengan harga yang kompetitif. Sedangkan kepemimpinan produk yang mereka maksudkan adalah perusahaan harus menyediakan produk secara terus-menerus, dan untuk keintiman pelanggan, yang dimaksud oleh mereka adalah menjual kepada pelanggan solusi total bukan hanya produk saja atau layanan saja. Walaupun ketiga strategi tersebut dapat mendisplinkan suatu perusahaan namun ketiga hal ini tidak boleh berjalan masing-masing karena setiap perusahaan memiliki kekuatan yang berbeda [2].

Dalam mengolah relasi dengan pelanggan, pesan yang nantinya akan dikelola bukanlah secara komunikasi tatap muka atau pada saat bertemu saja. Namun, mengolah relasi dengan pelanggan juga dapat menggunakan komunikasi secara *online* atau dengan melihat dan mendengarkan pengalaman mereka saat menggunakan produk atau jasa sebuah perusahaan[2]. Karena dalam mencari hotel saat ini wisatawan semakin mengandalkan *online review* daripada iklan atau promosi yang terdapat pada situs-situs *online*. Wisatawan lebih memilih membaca pengalaman yang telah dialami oleh pelanggan lain, sehingga akan lebih selaras dengan minat mereka sendiri dan untuk mengetahui apa yang diharapkan ketika mereka mencapai tujuan mereka. Jika tidak dipandu oleh pengalaman sebelumnya dan peringkat pelanggan, lalu memilih hotel baru pada saat pergi ke destinasi wisata biasanya merupakan tugas yang menantang bagi para wisatawan[3].

Rekomendasi secara langsung mungkin merupakan petunjuk penting untuk menjelaskan kepuasan dari pelanggan sebelumnya pernah menggunakan jasa tersebut hal ini dikemukakan oleh Xiang, Schwartz, Gerdes, dan Uysal (2015)[4]. Namun, belum ada penelitian yang menyelidiki secara terperinci apa yang mendorong rekomendasi langsung tersebut. Ketika pengalaman positif, perkataan langsung tersebut memiliki efek ajakan yang kuat pada pelanggan lain dengan menunjukkan bahwa pelanggan tidak hanya menikmati pengalaman tetapi juga memiliki keahlian dari staf yang tinggi. Namun,

rekomendasi yang negatif biasanya menyebabkan penurunan pembelian yang cukup besar[5].

Selama tahun terakhir, jumlah informasi yang tersedia secara *online* meningkat secara angka dan masalah kelebihan informasi telah menjadi salah satu tantangan utama yang dihadapi oleh pencarian informasi dan sistem penyaringan informasi. Sistem pemberi rekomendasi adalah salah satu solusi untuk masalah kelebihan informasi. Sistem pemberi rekomendasi dapat mengatasi masalah penyaringan informasi yang mungkin menarik bagi pengguna individu. Selain itu, memberikan rekomendasi kepada pengguna dengan mencerminkan selera pribadi mereka dan meyakinkan para pengguna untuk mempercayai dan mengeksplorasi rekomendasi yang diberikan adalah tujuan utama dari seorang pemberi rekomendasi. Aplikasi yang sukses telah digunakan di Internet oleh situs web pemberi rekomendasi hotel yang baik adalah *TripAdvisor.com*, yang merekomendasikan puluhan bahkan sampai ratusan hotel kepada calon pelanggan, pelanggan yang pernah datang ke hotel yang ada di *TripAdvisor.com* akan memberikan kesan dan pesan untuk hotel yang pernah ditinggali. Dari komentar atau *review* yang berikan pelanggan yang belum pernah berkunjung ke hotel tersebut dapat melihat apakah hotel itu bagus untuk di datangi atau tidak[6].

Analisis sentiment merupakan sebuah teknologi komputerisasi yang dapat membantu dan menganalisis sebuah kalimat opini tekstual yang cara bekerjanya adalah memahami dan mengekstraknya seperti halnya text mining sehingga dapat menghasilkan sebuah informasi sentimen. Analisis sentimen juga dilakukan untuk melihat kecenderungan seseorang berpandangan apakah beropini positif atau negatif, bahkan opini netral juga dapat membantu mendukung sebuah keputusan[7]. Sedangkan text mining adalah teknik atau cara untuk mengekstraksi informasi yang berasal dari data teks yang tidak terstruktur dan berguna[8].

Berdasarkan masalah yang didapatkan, penelitian ini akan berfokus untuk memberikan rekomendasi hotel untuk para wisatawan dari *online review* pada situs wisata terbesar yaitu *TripAdvisor.com*. Data *review* yang telah di

dapatkan akan di proses lalu akan mendapatkan hasil berupa ukuran sentimen pelanggan hotel. Hasil dari ukuran sentimen tersebut berupa pengkategorisasian *review* yang terbagi menjadi tiga kategori yaitu positif, negatif, dan netral. Ketiga kategori tersebut akan diproses lagi melalui dua tahap. Tahap yang pertama yaitu tahap menentukan rekomendasi *eksplisit* dan nantinya akan menghasilkan enam kategori yaitu positif rekomendasi, positif non rekomendasi, negatif rekomendasi, negatif non rekomendasi, netral rekomendasi, dan netral non rekomendasi. Sedangkan tahap kedua yaitu tahap menentukan rekomendasi *implisit* dan nantinya juga akan menghasilkan sepuluh kategori yaitu fully positif rekomendasi, fully positif non rekomendasi, semi positif rekomendasi, semi positif non rekomendasi, fully negatif rekomendasi, fully negatif non rekomendasi, semi negatif rekomendasi, semi negatif non rekomendasi, netral rekomendasi, dan netral non rekomendasi. *Review* dari masing-masing kategori nantinya akan diproses lagi dan akan menghasilkan beberapa kata yang sering muncul, kata-kata yang sering muncul itu akan di bagi menjadi 8 mode layanan. Penelitian ini berbasis kamus, untuk itu dalam menentukan rekomendasi dan 8 mode layanan akan dibuat dalam bentuk kamus yang akan digunakan dalam mencari hotel yang dapat direkomendasikan secara *eksplisit* maupun secara *implisit* dan juga menghasilkan mode layanan tambahan seperti apa yang dimiliki oleh hotel sehingga dapat direkomendasikan oleh calon pelanggan. Jika hotel tersebut hanya memiliki sebagian dari mode yang ada, maka akan dikembalikan kepada hotel agar dapat diperbaharui mode layanannya.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah disusun, permasalahan yang akan diselesaikan yaitu bagaimana menerapkan *text mining* dan Analisis Sentimen untuk merekomendasi hotel berdasarkan *online review* pada situs *TripAdvisor.com*?

1.3. Batasan Masalah

Berdasarkan latar belakang yang telah disusun agar penyelesaian tugas ini tidak menyimpang maka batas masalahnya adalah sebagai berikut .

- a. Data yang digunakan menggunakan *online review* dan data diambil dari situs *TripAdvisor.com*
- b. Data yang diambil menggunakan Bahasa Indonesia.
- c. *Online review* yang digunakan adalah *online review* dari tiga hotel dengan *review* terbanyak yang ada di Yogyakarta.

1.4. Tujuan Penelitian

Berdasarkan latar belakang yang telah disusun tujuan yang akan dicapai yaitu menerapkan *text mining* dan Analisis Sentimen untuk merekomendasi hotel berdasarkan *online review* pada situs *TripAdvisor.com*.

1.5. Metode Penelitian

Data yang digunakan dalam penelitian ini mengambil *online review* dari situs *TripAdvisor.com*. *Review* yang diambil adalah tiga *review* terbanyak saat pengambilan data dilaksanakan.

Metode yang akan digunakan dalam penelitian ini adalah sebagai berikut.

1. Melakukan pengumpulan data

Data yang diambil adalah text yang merupakan *online review* hotel yang diambil dari situs *TripAdvisor.com*.

2. Melakukan *pre-processing* data

Pada *pre-processing* data yang telah diambil tadi akan dibersihkan dengan menghilangkan beberapa *link*, tanda baca, *emoticon*, dan angka. Selain itu juga dilakukan pembersihan dengan menghilangkan *stopwords*, mengubah semua huruf menjadi *lowercase*, dan normalisasi kata.

3. Melakukan analisis dengan *text mining* dengan metode *Sentistrength*

Analisis ini dilakukan untuk menentukan *review* yang positif, negative, dan netral serta menggunakan metode *Customer Realitionship*

Management untuk mengetahui bagaimana prospek penjualan yang baik dengan menggabungkan pekerjaan manusia dan teknologi yang ada saat ini.

4. Melakukan analisis dengan *text mining* dengan menggunakan kamus rekomendasi

Analisis ini digunakan untuk mendapatkan hasil berupa jumlah pelanggan yang akan merekomendasikan suatu hotel melalui *review* yang telah ditulisnya pada situs *TripAdvisor.com*.

5. Membuat kesimpulan

Dalam hal ini akan berisikan mode layanan mana yang masuk dalam tambahan hotel tersebut sehingga dapat memperkuat rekomendasi dan mode layanan mana yang belum ada pada hotel yang nanti diharapkan dapat dikembangkan oleh pihak hotel dengan menggunakan prinsip kekuatan kompetitif.

1.6. Sistematika Penulisan

BAB I PENDAHULUAN

Pada bab ini berisikan konsep dasar dari pembuatan dan penulisan skripsi yaitu berupa latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini berisikan beberapa penelitian yang telah terlaksana yang dapat digunakan sebagai acuan untuk penelitian yang dilakukan. Di dalam bab ini juga terdapat tabel pembandingan yang membandingkan penelitian yang telah ada sebelumnya dengan penelitian yang dilakukan.

BAB III LANDASAN TEORI

Pada bab ini berisikan penjelasan berupa teori-teori yang membantu penulis dalam melaksanakan penelitian.

BAB IV ANALISIS DAN PERANCANGAN EKSPERIMEN

Pada bab ini berisikan deskripsi problem, analisis kebutuhan eksperimen, dan perancangan eksperimen.

BAB V HASIL EKSPERIMEN DAN PEMBAHASAN

Pada bab ini berisikan berisikan deskripsi eksperimen, hasil eksperimen, dan pembahasan eksperimen.

BAB VI PENUTUP

Pada bab ini berisikan kesimpulan dan saran yang didapatkan penulis dari penelitian yang telah dijalankan.



BAB II. TINJAUAN PUSTAKA

Review online umumnya bebas dari aturan yang memeriksa topik wajib dan pengulas dapat memilih untuk menulis dengan cara yang paling mewakili pengalaman mereka. Peneliti bebas untuk mengambil data negatif atau positif ketika mereka mendiskusikan atribut layanan, seperti perasaan mereka tentang pengalaman, sikap staf dan berapa banyak waktu yang mereka tunggu untuk dilayani. Namun, beberapa *review* secara *eksplisit* merekomendasikan layanan kepada pembaca, meskipun kalimat seperti "Saya sangat merekomendasikan restoran ini" atau "Saya pasti akan kembali" kadang-kadang dimasukkan dalam *review* untuk menunjukkan kepada rekan-rekan bahwa layanan ini adalah pengalaman yang layak. Pengesahan tersebut baru-baru ini dipelajari dalam literatur sebagai pendorong peringkat bintang yang dilaporkan sendiri dan niat beli. Meskipun tidak ada perjanjian yang jelas mengenai apakah rekomendasi langsung adalah *implisit* atau ekspresi *eksplisit*, ada kesepakatan bahwa itu adalah tindakan *direktif* yang sangat mempengaruhi keputusan konsumen. Namun, pengetahuan dan kesadaran konsumen diketahui mempengaruhi pengetahuan *eksplisit*, sedemikian rupa sehingga pengetahuan tersebut tercermin pada bagaimana pengulas menilai pengalaman mereka.

Memahami motivasi di balik rekomendasi *eksplisit* dan rekomendasi *implisit* dapat membantu manajer secara selektif meningkatkan penawaran mereka untuk memenuhi harapan konsumen dengan lebih baik. Namun demikian, karena sifat *review online* yang tidak terstruktur, sulit untuk membangun model yang menggunakan *text mining* untuk memprediksi rekomendasi *eksplisit* dan rekomendasi *implisit* tanpa menggunakan metode formal untuk menyusun informasi tersebut. *Text mining* telah berhasil digunakan dalam penelitian terbaru untuk mengatasi tantangan tersebut, dan metode ini digunakan dalam penelitian saat ini.

Pada tahun 2015, Lopes, Prajyoti, and Bidisha Roy melakukan penelitian untuk membuat sistem rekomendasi berdasarkan data navigational dari website menggunakan metode web mining. Penulis menggumpulkan *user click stream*

data dari sebuah website e-commerce. Data mentah tersebut dilakukan preprocessing berupa *field separation*, data bersih, *user differentiation*, *session identification*, *session clustering*, dan data format. Simpulan dari penelitian ini adalah sistem rekomendasi dapat memberikan hasil yang baik berdasarkan pola pengguna. Hasil dari sistem mengurangi kesalahan klasik dari sistem rekomendasi, yaitu pengurangan hasil positif yang salah[9].

Penelitian yang dilakukan oleh Adeniyi DA, Wei Z, dan Yongquan Y [10] bertujuan untuk membuat sistem rekomendasi berita menggunakan data yang didapatkan dari *Really Simple Syndication* (RSS). Algoritma klasifikasi KNearest Neighbor dilatih untuk dapat melihat pola klik dari pengguna, kemudian mengelompokkannya ke sebuah user group, kemudian memberikan rekomendasi informasi browsing sesuai dengan karakteristik pengguna. Proses penelitian dilakukan dengan data RSS di-ekstraksi, lalu dibersihkan datanya, diformat, dan dikelompokkan sehingga membentuk sebuah dataset yang dapat digunakan. Untuk menghitung jarak antar node, digunakan metode euclidean distance. Hasil dari penelitian ini, implementasi algoritma K-Nearest Neighbor dapat memberikan manfaat dan memberikan akurasi yang baik, sehingga rekomendasi yang dihasilkan sesuai dengan kebutuhan atau minat pengguna.

Penelitian terdahulu yang pernah dilakukan oleh D. Sebastian pada tahun 2017 adalah “Rancang Bangun Website Klasifikasi untuk Pencarian Produk Pasar *Online* menggunakan Algoritma KNearest Neighbor”[11]. Pada penelitian tersebut dilakukan simulasi perhitungan menggunakan data produk handphone. Hasil dari penelitian tersebut adalah pembuktian secara simulasi bahwa metode K-nearest neighbour dapat digunakan untuk kasus tersebut dan perancangan yang digunakan untuk membuat aplikasi website pada penelitian aktual.

Penelitian menggunakan algoritma K-Nearest Neighbor juga dilakukan oleh W. E. Nurjanah, R. S. Perdana dan M. A. Fauzi pada tahun 2017[12]. Penulis mengumpulkan data tweet dari media sosial twitter, dimulai dari tahapan preprocessing. Pada tahapan *pre-processing*, dilakukan *tokenisasi*, pembersihan data, *case folding*, *filterisasi*, dan *stemming*. Tahapan selanjutnya adalah pembobotan menggunakan metode TF-IDF. Hasil pembobotan dihitung jarak

kemiripannya menggunakan *cosine similarity*, kemudian diolah menggunakan algoritma K-Nearest Neighbor. Tahap selanjutnya adalah melakukan pembobotan jumlah Retweet dan dinormalisasi menggunakan metode *min-max*. Kemudian hasil penggabungan tersebut menghasilkan sebuah nilai, yang menunjukkan hasil klasifikasi dokumen positif atau negatif.

Pada tahun 2018, P. D. Nugraha, S. A. Faraby and Adiwijaya melakukan penelitian klasifikasi dokumen dan dikombinasikan dengan metode Information Gain[13]. Penelitian tersebut dilakukan dengan cara mencari atribut dari dataset, kemudian melakukan seleksi fitur menggunakan metode information gain kemudian diolah menggunakan klasifikasi K-Nearest Neighbor. Evaluasi dilakukan dengan cara membandingkan metode K-Nearest Neighbor dan kombinasi K-Nearest Neighbor dengan information gain. Kedua skenario tersebut dibandingkan waktu pemrosesan dan akurasi hasil pemrosesan. Berdasarkan tinjauan pustaka yang sudah dilakukan, maka diperoleh kesimpulan bahwa algoritma klasifikasi KNearest Neighbor dapat digunakan untuk mengelompokkan produk dari beberapa e-marketplace. Oleh karena itu, penelitian aktual dilakukan menggunakan metode pembobotan TF-IDF dan algoritma klasifikasi K-Nearest Neighbor.

Pada tahun 2019, Tang, F, et al melakukan penelitian *review online*. Penelitian tersebut dilakukan dengan menggunakan metode *Joint Aspect Based Sentiment Topic* (JABST). Hal yang dilakukan oleh penulis adalah bagaimana cara mengekstraksi secara efisien aspek multi-grained, mengidentifikasi opini yang terdapat dari user, dan mengklasifikasikan pola dari sentimen. Penulis mengusulkan untuk meningkatkan akurasi dan kinerja dengan menggunakan model JABST berbasis entropi (MacEnt-JABST). Hasil dari evaluasi dari penelitian adalah dengan adanya model JABST dan MAXEnt-JABST dapat mengungguli model yang lainnya[14].

Pada tahun 2018, Xiaolin Li, et al melakukan penelitian dampak bisnis dari *review online*. Penulis menggunakan metode *Sentiment-Topic Model*. Penelitian ini dilakukan karena adanya *review* yang dapat mempengaruhi hasil dari kinerja atau penjualan produk. Penulis berasumsi bahwa dengan adanya

peringkat secara numerik lebih berefek baik dibandingkan dengan sentimen secara tekstual. Temuan yang dilakukan penulis tidak hanya berkontribusi pada pengetahuan *review online* yang dapat berdampak pada penjualan produk, namun juga menggambarkan bagaimana peringkat *review* yang berbasis numerik dan tekstual saling mempengaruhi penjualan produk. Hasil dari penelitian ini agar dapat membantu vendor membuat strategi yang baik dengan berfokus pada aspek yang lebih relevan yang dapat mendorong penjualan[15].

Pada tahun 2019, Zhu, L, et al melakukan penelitian dengan tujuan untuk memecahkan kode kepuasan pelanggan dengan menganalisis sentimen pelanggan dengan peringkat *online*. Penulis menggunakan metode sentiment analisis, model Tobit dan Ordered Logit. Berdasarkan hasil penelitian penulis menemukan hal positif sentiment atau hal negatif sentiment dapat dikaitkan dengan peringkat tinggi atau rendahnya pariwisata atau hotel. Penulis menunjukkan bahwa tautan ini lebih kuat ketika pelanggan mewujudkan pemikiran analitis dan keaslian yang lebih tinggi. Model Tobit dan model logit yang dipesan menghasilkan hasil estimasi yang konsisten dari model lainnya[16].

Pada tahun 2017, Amplayo, R. K., et al melakukan penelitian dengan mengkaji metode baru dalam menghasilkan rangkuman dari beberapa *review online*. Metode yang digunakan penulis dalam melakukan penelitian ini dibagi menjadi 2 metode. Metode pertama adalah *Extended Biterm Topic Model* (eBTM) yang digunakan untuk mengklasifikasikan rangkuman dari teks yang sangat panjang. Metode kedua menggunakan LDA untuk mengrangkum text yang singkat atau text yang pendek. Secara keseluruhan, hasil menunjukkan bahwa pengklasifikasian sentiment lebih unggul dari pada model yang dasar dan pengklasifikasian standar industri lebih unggul dari model topik lainnya dalam hal keragaman aspek dan kekuatan ekstraksi aspek[17].

Pada tahun 2018, Ireland, R., et al melakukan penelitian dengan membuat kerangka kerja desain untuk menganalisis *review* produk *online*. Tujuannya adalah untuk menggunakan data yang dihasilkan mesin ini untuk mengidentifikasi serangkaian kebutuhan pelanggan. Kerangka kerja ini bertujuan untuk menyaring volume besar data kualitatif menjadi wawasan kuantitatif tentang fitur produk,

sehingga desainer dapat membuat keputusan yang lebih tepat. Itu kerangka kerja menggabungkan unsur-unsur *review* produk *online*, teori desain dan metodologi, dan analisis data untuk mengungkapkan wawasan baru. Efektivitas kerangka proposal divalidasi melalui studi kasus tentang tinjauan produk dari situs web e-commerce, Amazon. Kerangka kerja menunjukkan pendekatan statistik untuk menganalisis *review* produk *online*. Kerangka kerja ini bertindak sebagai antarmuka antara output kuantitatif dan proses desain kualitatif dan kreatif. Analisis lebih lanjut hasil mengidentifikasi banyak menggabungkan metode logis, komputasi ke dalam proses desain yang subyektif dan kreatif[18].

Pada tahun 2019, Kumar, A., et al melakukan penelitian dengan menawarkan model pembelajaran dalam hybrid untuk berbutir halus prediksi sentimen dalam data multimodal waktu nyata. Ini memperkuat kekuatan jaring pembelajaran mendalam dalam kombinasi dengan pembelajaran mesin untuk menangani dua sistem semiotik tertentu, yaitu teks (teks tertulis) dan visual (gambar diam) dan kombinasinya dalam konten *online* menggunakan tingkat keputusan penggabungan multimoda. Model ConvNet-SVMBoVW kontekstual yang diusulkan, memiliki empat modul, yaitu, diskritisasi, analisis teks, analisis gambar, dan modul keputusan. Modul analisis teks menentukan sentimen menggunakan hibrida dari jaringan saraf konvolusi (ConvNet) yang diperkaya dengan semantik kontekstual SentiCircle. Modul keputusan Boolean dengan operasi OR logis ditambahkan ke arsitektur yang memvalidasi dan mengkategorikan output berdasarkan lima kategori sentimen berbutir halus (nilai kebenaran), yaitu 'sangat positif,' 'positif,' 'netral,' 'negatif' dan 'sangat negatif'[19].

Pada tahun 2019, Kincl., et al melakukan penelitian dengan menyajikan model analisis sentimen bahasa dan domain-independen berdasarkan karakter n-gram yang meningkatkan kinerja pengklasifikasi dengan memanfaatkan konteks sekitarnya. Empat percobaan pada berbagai dataset dilakukan untuk memvalidasi model. Dataset termasuk korpus referensi yang berisi *review* film dalam bahasa Inggris, *review* film dalam bahasa Ceko, buku terlaris Amazon 2012 Fifty Shades of Grey *review* dataset dataset dari tiga mutasi bahasa Amazon (Inggris, Jerman, dan Perancis), corpus referensi lain yang mengandung *review* dari situs Amazon

dalam berbagai bahasa (Jerman, Prancis, dan Jepang), dan dataset multi-domain (film, buku, dan kategori produk mulai dari elektronik dan peralatan rumah tangga hingga perlengkapan olahraga dan perlengkapan untuk hobi dan hewan peliharaan). Eksperimen mengkonfirmasi pendekatan menggabungkan konteks sekitarnya agar efektif untuk dataset dari berbagai bahasa dan domain, menunjukkan kinerja yang kuat dari model berbasis karakter n-gram untuk multi-domain dan dataset bahasa juga. Klasifikasi all-in-one sederhana, yang menggunakan campuran data berlabel dari berbagai bahasa (atau domain) untuk melatih model klasifikasi sentimen, dapat menyaingi teknik adaptasi domain / bahasa yang lebih canggih. Pendekatan semacam itu mencerminkan kebutuhan perusahaan dengan keterkaitan dunia saat ini, sebagian besar perusahaan beroperasi di berbagai pasar dan akan kesulitan mendapatkan solusi analisis sentimen khusus untuk setiap pasar yang mereka layani[20].

Pada tahun 2019, Atzeni, M., et al melakukan penelitian menyajikan pendekatan mutakhir untuk klasifikasi polaritas sentimen. Penelitian bergantung pada ansambel jaringan Memori Jangka Pendek Panjang Dua Arah yang dilengkapi dengan mekanisme perhatian saraf. Sistem ini menggunakan kata latih dan mampu memprediksi vektor baru untuk kosa kata, dengan mempelajari representasi distribusi berdasarkan ejaan kata. Selain itu, selama proses pelatihan jaringan saraf berulang digunakan untuk melakukan *fine-tuning* dari kata asli, dengan mempertimbangkan informasi akun tentang polaritas sentimen. Langkah ini dapat sangat membantu untuk analisis sentimen, karena kata biasanya dibangun berdasarkan informasi konteks, sedangkan kata-kata dengan polaritas sentimen yang berlawanan sering terjadi dalam konteks yang sama. Sistem yang dijelaskan dalam makalah ini adalah versi perbaikan dari pendekatan yang bersaing dalam tantangan baru pada analisis sentimen semantik. Penulis mengevaluasi kinerja sistem pada set uji *multi-domain* yang sama yang digunakan oleh penyelenggara tantangan, hal ini menunjukkan bahwa pendekatan tersebut memungkinkan mencapai hasil yang lebih baik sehubungan dengan sistem penilaian atas sebelumnya. Lalu, penulis menanamkan pendekatan polaritas sentimen yang diusulkan di atas robot *humanoid* untuk secara aktif

mengidentifikasi sentimen pengguna yang berbicara[21].

Pada tahun 2017, Lee, J. H., et al melakukan penelitian dengan berfokus pada efek *Word of Mouth* (WOM) secara *online* dengan mengukur volume dan valensi WOM. Namun, beberapa penelitian telah menyelidiki bagaimana entropi dari sentimen teks *review* mempengaruhi hubungan antara WOM *online* dan penjualan produk. Karena valensi dan volume WOM biasanya diberikan pada tingkat agregat, konsumen seringkali tidak memiliki informasi yang cukup untuk membuat keputusan. Dalam hal ini, membaca teks *review online* telah menjadi proses penting bagi konsumen untuk membuat keputusan pembelian. Namun, ketika pelanggan menjumpai terlalu banyak teks *review* positif, mereka mungkin meragukan kredibilitas WOM *online* itu sendiri. Penulis menganalisis entropi dari sentimen teks *review* dengan melakukan teknik penambangan teks. Penulis mengklasifikasikan sentimen teks *review* ke dalam kategori positif, negatif, dan netral dan menciptakan variabel entropi. Entropi tingkat tinggi dalam teks *review* menunjukkan bahwa sentimen dari teks *review* didistribusikan secara merata, tetapi tidak bias, terhadap sentimen positif atau negatif[22].

Pada tahun 2019, Chongchao, C., et al melakukan penelitian untuk mengatasi kualitas prediksi rekomendasi teman di jejaring sosial, algoritma rekomendasi pengguna untuk jejaring sosial berdasarkan analisis sentimen dan faktorisasi matriks. Metode ini didasarkan pada model faktorisasi matriks tradisional. Dengan mengintegrasikan Sentimen (S), Penting (I) dan Tujuan (O) dari konten topik pengguna di jejaring sosial, penulis mengusulkan basis pendekatan pada analisis sentimen dan faktorisasi matriks untuk menyelesaikan akurasi prediksi yang buruk dengan menggunakan jejaring sosial. Model SIO memecahkan masalah yang membuat pengguna di jejaring sosial tidak dapat menilai konten topik. Matriks Usertopik dibangun oleh model SIO. Menggabungkan model SIO dengan faktorisasi matriks, algoritma yang disebut algoritma SIO-TMF diusulkan. Menerapkan metode ini di jejaring sosial, membandingkan dengan beberapa algoritma rekomendasi tradisional dari empat aspek: akurasi, keragaman, kebaruan dan cakupan, hasil percobaan menunjukkan bahwa metode yang diusulkan meningkatkan kualitas prediksi sistem

merekomendasikan[23].

Perbedaan penelitian aktual dengan penelitian terdahulu dapat dilihat pada
TABEL 2.1.



Tabel 2.1. Perbandingan Penelitian

Judul	Dynamic Recommendation System Using Web Usage Mining for E-commerce Users	Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method	Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode KNearest Neighbor dan Pembobotan Jumlah Retweet	Rancang Bangun Website Klasifikasi untuk Pencarian Produk Pasar <i>Online</i> menggunakan Algoritma K-Nearest Neighbor	Implementasi Algoritma K-Nearest Neighbor Untuk Melakukan Klasifikasi Produk dari Beberapa E-Marketplace.	Klasifikasi Dokumen Menggunakan Metode Knearest neighbour (KNN) dengan Information Gain.	Penerapan <i>text mining</i> dan Analisis Sentimen untuk merekomendasi hotel berdasarkan <i>online review</i>
-------	---------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------

Oleh	P. Lopes and B. Roy	Adeniyi DA, Wei Z, Yongquan Y	W. E. Nurjanah, R. S. Perdana and M. A. Fauzi	Danny Sebastian	Danny Sebastian	P. D. Nugraha, S. A. Faraby and Adiwijaya	Margareta Oktaviana Shintapuri
Metode	User Interest Measure, Wish list buffer, Similarity Measures	Eucledian distance, KNearest Neighbor	KNearest Neighbor dengan pembobotan jumlah retweet menggunakan min-max	TF-IDF, Eucledian distance, dan KNearest Neighbor	TF-IDF, Eucledian distance, dan KNearest Neighbor	KNearest Neighbor dan Information Gain	Sentistength
Object	Navigational data di website	Really Simple Syndication (RSS)	Tweet dari media sosial twitter	Produk dari tokopedia dan bukalapak	Produk dari tokopedia dan bukalapak	Text Categorization Collection Data Set	<i>Review Online</i> dari situs <i>TripAdvisor.com</i>
Tahun	2015	2016	2017	2017	2018	2018	2020

BAB VI. PENUTUP

6.1. Kesimpulan

Berdasarkan penelitian penambangan teks atau *text mining* dengan menggunakan metode *sentistrength* yang dilakukan dengan menggunakan data tiga *review* terbanyak pada hotel yang ada di D.I. Yogyakarta dan dari situs *TripAdvisor.com* yaitu Hotel Eastparc, Hotel The Phoenix, dan Hotel Hotel Tentrem. Penulis dapat menyimpulkan bahwa ketiga hotel tersebut direkomendasikan secara *eksplisit* oleh sedikit pelanggan yang pernah menggunakan jasa ketiga hotel tersebut. Namun, ketiga hotel tersebut direkomendasikan oleh banyak pelanggan secara *implisit* melalui komentar pelanggan tersebut yang memuji terus menerus jasa yang telah pelanggan terima.

Ketiga hotel tersebut dapat direkomendasikan secara *eksplisit* ataupun *implisit* dikarenakan layanan yang diberikan kepada hotel tersebut. Tiga layanan tambahan hotel yaitu elemen layanan perhotelan, elemen informasi, dan elemen konsultasi merupakan elemen utama yang dapat memperkuat rekomendasi sedangkan lima layanan tambahan lainnya yaitu elemen penagihan, elemen pemberian layanan pemesanan, elemen pembayaran, elemen menjaga keamanan, dan elemen pengecualian merupakan faktor pendukung dari direkomendasikannya jasa atau layanan. Pihak hotel dapat menggunakan penelitian ini untuk mengembangkan dan memperbaiki layanan yang telah ada pada hotel agar menambah presentase direkomendasikannya hotel tersebut.

6.2. Saran

Setelah dilakukannya penelitian, penulis memiliki beberapa saran untuk penelitian berikutnya yang juga akan melakukan analisis rekomendasi melalui *review online* yang menggunakan metode *sentistrength*, yaitu sebagai berikut :

1. Analisis rekomendasi pada data berupa teks dari *review online* dapat dilakukan pada situs dan media sosial lain selain *TripAdvisor.com*.

2. Penelitian untuk menganalisis rekomendasi juga dapat diterapkan pada perusahaan lain selain hotel.
3. Kamus rekomendasi secara *eksplisit*, kamus mode layanan, dan kamus sentimen dalam Bahasa Indonesia dapat ditambah lagi sehingga hasilnya lebih akurat lagi sehingga dapat sesuai dengan keinginan peneliti.
4. Pada tahap normalisasi kata di bagian pre-processing juga ditambahkan agar pada saat proses data, data yang dihasilkan benar-benar bersih.



DAFTAR PUSTAKA

- [1] D. P. D. I. Yogyakarta, “Statistik Kepariwisata 2018,” vol. 3, no. 2, pp. 54–67, 2018.
- [2] A. Payne, *Handbook of CRM: achieving excellence in customer management*, vol. 4, no. 3. 2012.
- [3] H. C. M. V. A. N. Trijp, *Journal of Consumer Research Advance Access published November 17, 2015* 1. 2015.
- [4] Z. Xiang, Z. Schwartz, J. H. Gerdes, and M. Uysal, “What can big data and text analytics tell us about hotel guest experience and satisfaction?,” *Int. J. Hosp. Manag.*, vol. 44, pp. 120–130, 2015.
- [5] G. Packard and J. Berger, “How language shapes word of mouth’s impact,” *J. Mark. Res.*, vol. 54, no. 4, pp. 572–588, 2017.
- [6] W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [7] N. Boudad, R. Faizi, R. Oulad, H. Thami, and R. Chiheb, “Sentiment analysis in Arabic : A review of the literature,” *Ain Shams Eng. J.*, 2017.
- [8] B. S. Kumar and V. Ravi, “A survey of the applications of text mining in financial domain,” *Knowledge-Based Syst.*, vol. 114, pp. 128–147, 2016.
- [9] P. Lopes and B. Roy, “Dynamic recommendation system Using web usage mining for E-commerce users,” *Procedia Comput. Sci.*, vol. 45, no. C, pp. 60–69, 2015.
- [10] D. A. Adeniyi, Z. Wei, and Y. Yongquan, “Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method,” *Appl. Comput. Informatics*, vol. 12, no. 1, pp. 90–108, 2016.
- [11] D. Sebastian, “Rancang Bangun Website Klasifikasi Untuk Pencarian Produk Pasar Online Menggunakan Algoritma K-Nearest Neighbor,” vol. 3, pp. 417–432, 2017.

- [12] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017.
- [13] P. D. Nugraha and S. Al Faraby, "Klasifikasi Dokumen Menggunakan Metode k -Nearest Neighbor (kNN) dengan Information Gain Document Classification using k- Nearest Neighbor (k NN) Method with Information Gain," vol. 5, no. 1, pp. 1541–1550, 2018.
- [14] F. Tang, L. Fu, B. Yao, and W. Xu, "Aspect based fine-grained sentiment analysis for *online reviews*," vol. 488, pp. 190–204, 2019.
- [15] X. Li, C. Wu, and F. Mai, SC. Elsevier B.V., 2018.
- [16] L. Zhu, Y. Lin, and M. Cheng, "International Journal of Hospitality Management Sentiment and guest satisfaction with peer-to-peer accommodation : When are *online* ratings more trustworthy?," *Int. J. Hosp. Manag.*, no. 3688, p. 102369, 2019.
- [17] R. K. Amplayo, M. Song, R. K. Amplayo, and M. Song, "Author ' s Accepted Manuscript *online reviews*," *Data Knowl. Eng.*, 2017.
- [18] R. Ireland and A. Liu, "CIRP Journal of Manufacturing Science and Technology Application of data analytics for product design : Sentiment analysis of *online* product reviews," *CIRP J. Manuf. Sci. Technol.*, 2018.
- [19] A. Kumar, K. Srinivasan, C. Wen-huang, and A. Y. Zomaya, "Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data," vol. 57, no. July 2019, 2020.
- [20] M. Nov, "Improving sentiment analysis performance on morphologically rich languages : Language and domain independent approach I a," vol. 56, pp. 36–51, 2019.
- [21] M. Atzeni and D. R. Recupero, "Multi-domain sentiment analysis with mimicked and polarized word embeddings for human – robot interaction," *Futur. Gener. Comput. Syst.*, no. xxxx, 2019.

- [22] J. H. Lee and J. Park, "The Role of Entropy of *Review* Text Sentiments on *Online* WOM and Movie Box Office Sales," *Electron. Commer. Res. Appl.*, 2017.
- [23] J. V. C. I. R, C. Cai, and H. Xu, "A topic sentiment based method for friend recommendation in *online* social networks via matrix factorization q," *J. Vis. Commun. Image Represent.*, vol. 65, p. 102657, 2019.
- [24] M. A. Hearst, "Text Data Mining," *Oxford Handb. Comput. Linguist.*, vol. 9780199276, 2012.
- [25] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDT) Data Structure : the Concept Hierarchy Concept Distributions," *Proc. 1st Int. Conf. Knowl. Discov. Data Min.*, pp. 112–117, 1995.
- [26] N. Katayama, O. Nohara, H. Moriyama, and H. Fujimaki, "Attempt to isolate mast-cell precursors based on the differential sensitivity to UV-B and X-irradiation," *Toxic Subst. J.*, vol. 13, no. 2, pp. 85–95, 1994.
- [27] P. Use, Q. Of, and B. Managers, "Data-mining basics."
- [28] M. Osman, A. Hilal, and M. Alhawarat, *Fine-Grained Quran Dataset*, vol. 6, no. 12. 2015.
- [29] G. Vinodhini and R. M. Chandrasekaran, "A comparative performance evaluation of neural network based approach for sentiment classification of *online reviews*," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 28, no. 1, pp. 2–12, 2016.
- [30] H. C. Soong, N. B. A. Jalil, R. Kumar Ayyasamy, and R. Akbar, "The essential of sentiment analysis and opinion mining in social media : IIntroduction and survey of the recent approaches and techniques," *ISCAIE 2019 - 2019 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 272–277, 2019.
- [31] D. Bandorski *et al.*, "Contraindications for video capsule endoscopy," *World J. Gastroenterol.*, vol. 22, no. 45, pp. 9898–9908, 2016.
- [32] K. Dave, I. Way, S. Lawrence, and D. M. Pennock, "P519-Dave," pp. 519–528, 2003.
- [33] M. Liu, L. Zhang, Y. Liu, H. Hu, and W. Fang, "Recognizing semantic correlation in image-text weibo via feature space mapping," *Comput. Vis.*

Image Underst., vol. 163, pp. 58–66, 2017.

- [34] X. Liu, “Full-Text Citation Analysis : A New Method to Enhance,” *J. Am. Soc. Inf. Sci. Technol.*, vol. 64, no. July, pp. 1852–1863, 2013.
- [35] M. R. Islam and M. F. Zibran, “SentiStrength-SE: Exploiting domain specificity for improved sentiment analysis in software engineering text,” *J. Syst. Softw.*, vol. 145, no. July 2017, pp. 125–146, 2018.

